

General Disclaimer

One or more of the Following Statements may affect this Document

- This document has been reproduced from the best copy furnished by the organizational source. It is being released in the interest of making available as much information as possible.
- This document may contain data, which exceeds the sheet parameters. It was furnished in this condition by the organizational source and is the best copy available.
- This document may contain tone-on-tone or color graphs, charts and/or pictures, which have been reproduced in black and white.
- This document is paginated as submitted by the original source.
- Portions of this document are not fully legible due to the historical nature of some of the material. However, it is the best reproduction available from the original submission.

Houston Univ. Texas.

NASA GRANT NSG-7440

PROGRESS REPORT 6/1/81-present

I. 5S rRNA Results

(A) Primary Sequence Characterizations

A major issue that was pursued in the past period was examination of the extent of phylogenetic distribution of the large 5S rRNA from Halococcus morrhuae. It was hoped in this way to determine if the insertion was of recent origin, i.e., very localized phylogenetically or representative of a more ancient event. We thus prepared ribonuclease T₁ primary fingerprints and conducted secondary and tertiary procedures necessary to characterize the 35 distinct strains of extremely halophilic bacteria listed in Table 1. In addition, pancreatic primary fingerprints were prepared for MS-3, Halobacterium marismortui, and Hb. sodomense.

These characterizations demonstrated that the large 5S rRNA is only found in the Halococcus strains, # 21-35 in Table 1. Whereas extensive 5S rRNA sequence variability was found in the Halobacterium strains, # 1-20, very limited variability was found among the Halococcus strains. While there is always one more strain that should have been considered, the evidence is very strong that the 5S rRNA insert is in fact localized phylogenetically. The most divergent large 5S rRNA was found in the Micrococcus strain (misnamed -- most assuredly a Halococcus strain), #31. This 5S rRNA differs in an estimated eleven positions out of a total of 231 but only three of these changes are known to be in the homologous region. Thus very little variability has been found in strains

(NASA-CR-174019) [PHYLOGENIC DISTRIBUTION
OF THE LARGE 5S rRNA FROM HALOCOCCUS
MORRHUAE] Progress Report (Houston Univ.)
30 p HC A03/MF A01

N85-11523

CSCl 06C

Unclass

G3/51 24897

TABLE 1: Strains Examined by Comparative Cataloging

<u>NAME</u>		<u>LARGE 5S rRNA</u>
1. <i>Halobacterium cutirubium</i>	NRC-34001	No
2. <i>Hb. volvani</i>	NCMB-2012	No
3. <i>Hb. species</i>	CCM-3361	No
4. <i>Hb. trpanicum</i>	NRC-34021	No
5. <i>Hb. sodomense</i>		No
6. <i>Hb. saccharovorum</i>	ATCC-29252	No
7. <i>Hb. saccharovorum</i>	NCMB-2081	No
8. <i>Hb. marismortui</i>	(Ginsburg)	No
9. <i>Hb. vallismortis</i>	NCMB-2082	No
10. <i>Hb. cutirubrum</i>	NCMB-763	No
11. <i>Hb. salinarium</i>	NCMB-786	No
12. <i>Hb. halobium</i>	NCMB-777	No
13. <i>Hb. trpanicum</i>	NCMB-784	No
14. <i>Hb. pharaonis</i>		No
15. <i>Ha. sinaiensis</i>		No
16. <i>Ha. californiae</i>		No
17. Alkaphilic strain	SP-1	No
18. Alkaphilic strain	SP-2	No
19. Alkaphilic strain	MS-3	No
20. Alkaphilic strain	SP-4	No
21. <i>Halococcus morrhuae</i>	ATCC 17082	Yes
22. <i>Hc. morrhuae</i>	NRC 16008	Yes

23. <i>Hc. morrhuae</i>	NRC 16012	Yes
24. <i>Hc. morrhuae</i>	NRC 16015	Yes
25. <i>Hc. morrhuae</i>	NRC 16016	Yes
26. <i>Hc. morrhuae</i>	NRC 16017	Yes
27. <i>Hc. morrhuae</i>	CCM-859	Yes
28. <i>Hc. morrhuae</i>	NCMB-757	Yes
29. <i>Hc. morrhuae</i>	ATCC 17077	Yes
30. <i>Hc. morrhuae</i>	NRC 16007	Yes
31. <i>Micrococcus</i> sp.	NRC 14043	Yes
32. <i>Hc.</i> species	NCMB 771	Yes
33. <i>Hb.</i> species	NCMB 708	Yes
34. <i>Bacterium halobium</i>	NCMB-738	Yes
35. unidentified sp.	NRC 41017	Yes

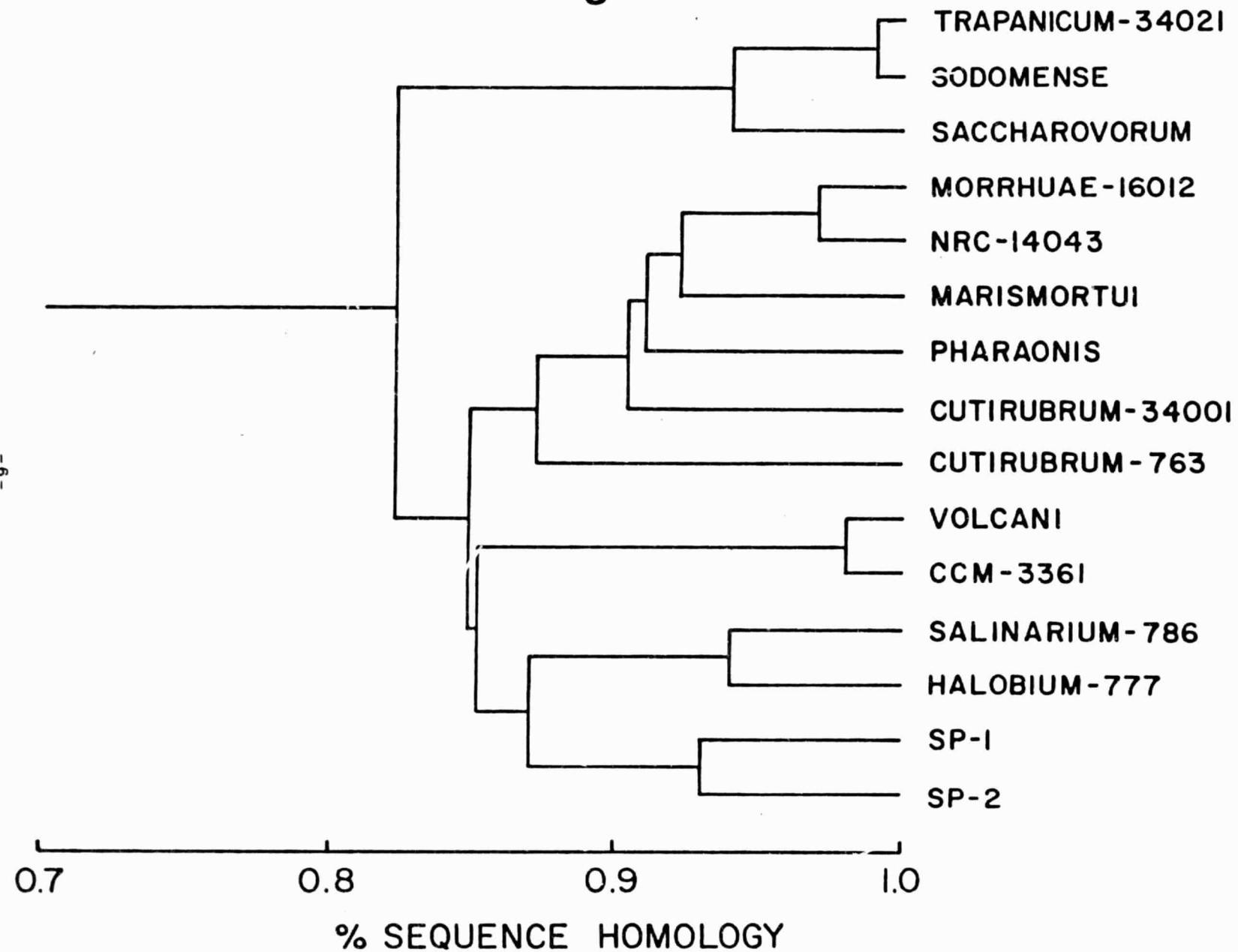
containing the large 5S rRNA which strongly suggests that the insertion was in fact of recent origin.

Because of the limited diversity seen in the strains containing the insertion, it has not yet been feasible to obtain the large amounts of comparative data needed to gain insight to the secondary structure of the insertion region. The catalog data does, however, provide a new awareness of the phylogenetic diversity among the extreme halophiles. Because complete sequences were already available for Halococcus morrhuae, Halobacterium volcani and Halobacterium cutirubium we are able to reliably place almost all of the ribonuclease T₁ oligomers. As a result one can infer partial sequences, Table 2, which cover approximately 77% of the molecule. Only 15 new strains are included in Table 2 because many of the strains, especially the Halococci give essentially identical fingerprints. The alignment allowed the calculation of a matrix of differences and construction of a tree, Figure 1. An examination of the location of sequence variation was made, Figure 2, and as in the case of other groups of closely related 5S rRNAs, certain "evolutionary hotspots" are apparent. Most notable among these are positions 105-111 since this is the region from which the Halococcus inseri emerges. Indeed, this is also the helix V region whose status in archaeabacteria is suspect.

Table 2

1	UUAG	GCCAUAG	GUUACUCCCCGUACCCAUCCCCGAACACCGGAAGAUAAAGCCC GCCUG	GUUCCGGGUUCAGUACUG	GCCUCUG GAAAUCGGGUUCGCCGCCUACU
2	PUUAG	GCCAG	GUUCCACCCGUACCCAUCCCCGAACACCGGAAGUUAGCUACCUG	GUUCUGGUUCAGUACUG	GAUCCUCUG GAAAUCAGUUCGCCGCCU
3	PUUACG	GCCAG	GUUCCACCCGUACCCAUCCCCGAACACCGGAAGUUAGCUCCUG	GUUCUGGUUCAGUACUG	GAUCCUCUG GAAAUCAGUUCGCCGCC
4	PUAAG	GCCAAAG	GCAACACCCGUACCCAUUCCGAACACCGGAAGUUAGCCACCAG	GUAUCCGGGAAGGUACUG	GAUCCUCUG GAACCAAG GUCGCCGCC
5	PUAAG	GCCAAAG	GAACACCCGUACCCAUUCCGAACACCGGAAGUUAGCCACCAG	GUAUCCGGGAAGGUACUG	GAUCCUCUG GAACCAAG GUCGCCGCC
6	PUAAG	GCCACAG	GACACCCGUACCCAUUCCGAACACCGGAAGUUAGCCACCAG	GUACCGGGGAAGGUACUG	GAUCCUCUG GAACAG GUUCGCCGCC
8	PUUAG	GCCACAG	GCCUCCCGUACCCAUUCCGAACACCGGAAGUUAGCCACCAG	GUUCCGUUG GUACUG	GAUCCUCUG GAAACG GUUCGCCGCC
9	PUUAG	GCCACAG	GCCUCCCGUACCCAUUCCGAACACCGGAAGAUAGCCCACCAG	GUUCCG GAAGUACUG	GCUCUG GAAACG GUUCGCCGCC
10	PUUAG	GCCACAG	GUACACCCGUACCCAUUCCGAACACCGGAAGUUAGCUGCCAC	GUUCG GUACUG	GCUCUG GAAACG GAAUCGCCGCC
11	PUUACG	GCCAUAG	GCCUCCLGUACCCAUUCCGAACACCGGAAGAUAGCUGCCUG	GUUUCGGGUUCAGUACUG	GACCCUCUG GAAAUCUGAUUCGCCGCC
12	PUUACG	GCCAUAG	GCCUCCCGUACCCAUUCCGAACACCGGAAGAUAGCUGCCUG	GUUACGGGUUAQUACUG	GACCCUCUG GAAUCUGGUUCGCCGCC
13	PUUACG	GCCAUAG	GCCUCCCGUACCCAUUCCGAACACCGGAAGAUAGCUCCUG	GUUACGGGUUAQUACUG	GACCCUCUG GAAUCUGGUUCGCCGCC
21	PUUAG	GCCACAG	GACUCCCGUACCCAUUCCGAACACGGCAGAUAGCCCACCAG	GUUCCAG GUACUG	GAACCUUC GAAAACUGGUUCGCCGCC
29	PUUAG	GCCACAG	GACUCCCGUACCCAUUCCGAACACGGCAGAUAGCCCACCAG	GUUCCAG GUACUG	GAACCUUC GAAAACUGGUUCGCCGCC
31	PUUAG	GCCACAG	GACUCCCGUACCCAUUCCGAACACGGCAGAUAGCCCACCAG	GUUCCG GUACUG	GCUCUG GAAAAAUUGGUUCGCCGCC
17	PUUAG	GCCAUUC	GUUCCUCCCGUACCCAUUCCGAACACGGGAAGAUAGCUCCUG	GUAUUG GUACUG	GACCCUCUG GCCGAUUCGCC
18	PUUAG	GCCAUUC	GUUCCUCCCGUACCCAUUCCGAACACGGGAAGAUAGCUCCUG	GUAUUG GUACUG	GACCCUCUG GCGGAUUCGCC
19	PUUAG	GCCAUUC	GUUCCUCCCGUACCCAUUCCGAACACGGGAAGAUAGCUCCUG	GUAUUG GUACUG	GACCCUCUG GCGGAUUCGCC
20	PUUAG	GCCAG	GUUCCUCCCGUACCCAUUCCGAACACGGGAAGAUAGCCCGCCCG	GUAUCCGGAAACUACUG	GACCCUCUG GAAACGUUCGUUCGCC
14	PUUAG	GCCACAG	GACCCUCGGGUACCCAUUCCGAACACGGGAAGAUAGCCCACCAG	GUUCCAG GUACUG	GCUCUG GAAAUCUGGUUCGCCGCC
16	PUUAG	GCCACAG	GCCUCCCGUACCCAUUCCGAACACGGGAAGAUAGCCCACCAG	GUUCCGUUG GUACUG	GCUCUG GAAAACG GUCGCCGCC
15	PUUAG	GCCACAG	GCCUCCCGUACCCAUUCCGAACACGGGAAGAUAGCCCACCAG	GUUCCGUUG GUACUG	GCUCUG GAAACG GUUCGCCGCC
36 ^{CG}		GACCAUCAG	GCACCAUCAGUUCCCAUUCCGAACACGGGAAGAUAGCCCACCAG	CCG AUGGUASUG	GAGCCCAUCGGAUAGUAGUSCAUCGUAGGCC

Figure 1



9:
 8:
 7:
 6:
 5:
 4:
 3:
 2:
 1:
 A: :::: A G x :: :: G C C A C A G x :: :: :: :: G U G C C U C C C G U A C C C A U C C C G
 B: UUA A G G C G G C C A U A G C G G U G G G U U A C U C C G U A C C C C A U C C C G
 10 20 30 40

10:
 9:
 8:
 7:
 6:
 5:
 4:
 3:
 2:
 1:
 A: A A C A C G S A A G A U A A G C C C G C C A G :: G U U C C G G U A A G U A C U G x
 B: A A C A C G G A A G A U A A G C C C G C C U G C G U U C C G G U C A G U A C C U G G
 50 60 70 80

10:
 9:
 8:
 7:
 6:
 5:
 4:
 3:
 2:
 1:
 A: :::: x x x x x G A G C C U C U G x G A A A U C C G G U U C C G G C C G H H H H H H
 B: A G U G C G C G A G C C U C U G G G A A A U C C G G U U C C G G C C G C C U A C U OH
 90 100 110 120

A: Master sequence for 17 halophilic bacteria.

B: *Bb. cutirubrum* NRC 34001 sequence numbers below are base position numbers.
 Symbols: Transversions, F=A-U, T=A-C, I=G-U, and Z=G-C; Transitions, C-G-A, T-C-U.

Figure 2

(B) 5S rRNA Complete Sequences

During the grant period a number of additional 5S rRNA eubacterial and eukaryotic sequences have been determined. These include: Pseudomonas cepacia, Ps. diminuta, Ps. maltophilia (2 strains), Ps. fluorescens, Ps. bathycetes, Ps. strain E-266, Rhodopseudomonas gelatinosa, Alteromonas, a barophile strain, Prosthecochloris, Pelodictyon, Clostridium Strain Howard 1, Treponema phagedensis, Treponema denticola, and Eimeria tennella. Partial sequences and sequences in progress include Leucosporidium scotti, Prototheca wickerham, Amoeba proteus, Polychaos dubia, Chaos chaos, and Desulfurococcus.

Our comparative sequencing work suffered a severe setback in the past grant period due to the untimely death of Dr. Siavash Baharaeen who as a post-doctoral research associate was involved in this area of the project. Two interesting eukaryotic sequences, Leucosporidium and Prototheca will be especially difficult to complete in his absence. Dr. Joanna Michalik of the Institute of Biochemistry and Biophysics in Warsaw, Poland has now taken over these efforts and sequences are currently being determined again at good rates.

The Pseudomonas 5S rRNA sequences have proven particularly useful in defining the major subdivisions in the purple group of the gram negative bacteria and were relied on by Stahl et al., 1984 for this purpose. The agreement with 16S rRNA data is in fact quite good.

The green photosynthetic bacteria and Treponemas are important because they provide the first representatives of two major eubacterial branches. In addition all of these organisms contain a very abbreviated helix IV, Figure 3 where in the case of Treponema three base pairs have been deleted relative to the usual eubacteria, e.g. E. coli. A similar phenomenon is seen in certain

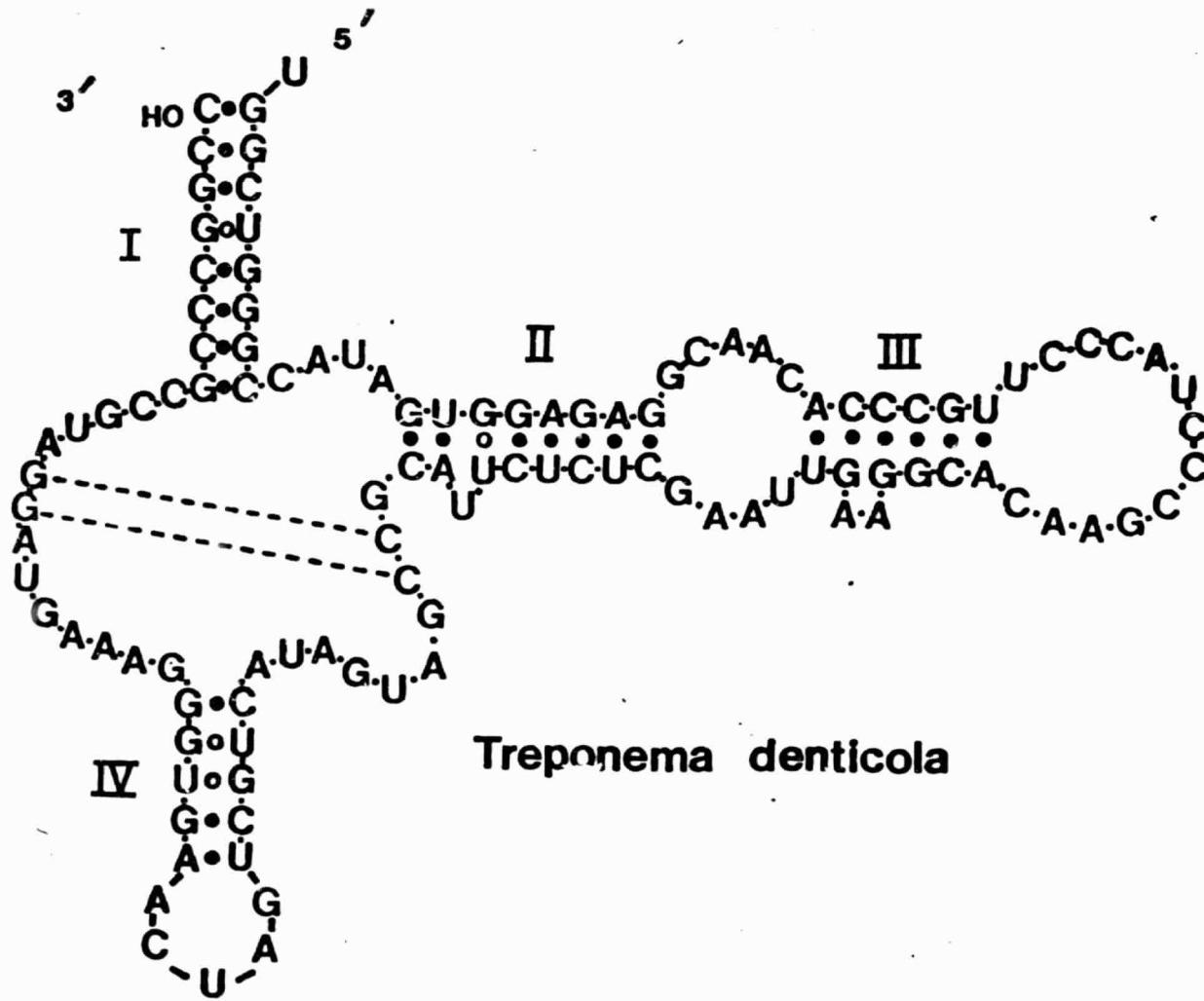


Figure 3: Treponema 5S rRNAs have an abbreviated Helix IV.

mycoplasma 5S rRNAs. This shortened helix is clearly an eubacterial structural variant. It is not yet clear if it is indicative of evolutionary convergence or a deep specific phylogenetic relation between the organism that have it.

The eukaryotic sequences, especially the amoeba's represent an important addition to the protozoan group. We have however encountered technical problems of an undefined nature that have made them difficult to sequence. The Clostridium strain will provide more substance to the gram positive group and the nearly complete Desulfurococcus sequence will provide representation of another major eubacterial branch in the 5S rRNA data base.

C. 5S rRNA Data Base

We have continued to maintain a current data base of 5S rRNA sequences. This has been increasingly difficult in the face of an explosion of information being published and we have had to significantly revise our procedures to effectively utilize the data. At present we can readily produce trees using Li's algorithm (PNAS 78, 1085-1089, 1981) and have routinely done so in the past year in order to follow the newly emerging results. Recently we implemented Fitch's improved method for tree construction (J. Mol. Evol. 18, 30-37, 1982). Among the most interesting new results in the 5S trees in the past year are those pertaining to the algae. In the case of some green algae there is a relationship to the higher plants, as expected. The brown algae that have been examined so far and Euglena cluster with the protozoans and are thus quite separated from the green algae. The red algae do, however, appear to be distantly related to the green algae. The position of the golden brown algae remains unknown. We feel that the algae require significant additional work and since they are being largely ignored by other workers in the 5S rRNA field we intend to expand our eukaryotic efforts in this direction in the next year.

We also have continued to examine the structural properties of the 5S rRNAs. Although there are structural variants among the eubacteria, most fit the consensus secondary structure shown in Figure 4 which clearly differs from that of the eukaryotes. The location of universal nucleotides, Figure 4, supports this too as it clearly differs in the two 5S rRNA types.

II. 16S rRNA Results

A. 16S rRNA Catalog Processing

Analysis of 16S rRNA catalog data is conducted as it is received from collaborating laboratories at the University of Illinois and the University of Kiel in Kiel, West Germany. This information is keypunched, proofed and a comparison made to a sample data set to check for format errors. The new catalog is then entered into a random access file and a listing printed. This listing is returned to the contributing investigator for further verification. Any errors, or catalog amendments, can be readily implemented through the use of an update program. Matching coefficients are calculated for all binary comparisons of 16S rRNA catalogs. Matrices of matching coefficients are constructed and dendograms generated by any of several clustering algorithms including average linkage, complete linkage or single linkage. When desired, consolidated tables of oligomers from numerous catalogs can be constructed. Also, any two catalogs can be compared in order to identify related oligonucleotides or the entire data set can be scanned to assess which organisms contain a particular oligonucleotide.

B. Phylogenetic Results Based on 16S rRNA

Recently we have completed generation of new summary dendograms for all 450 organisms for which catalog data or complete sequence data (which can be

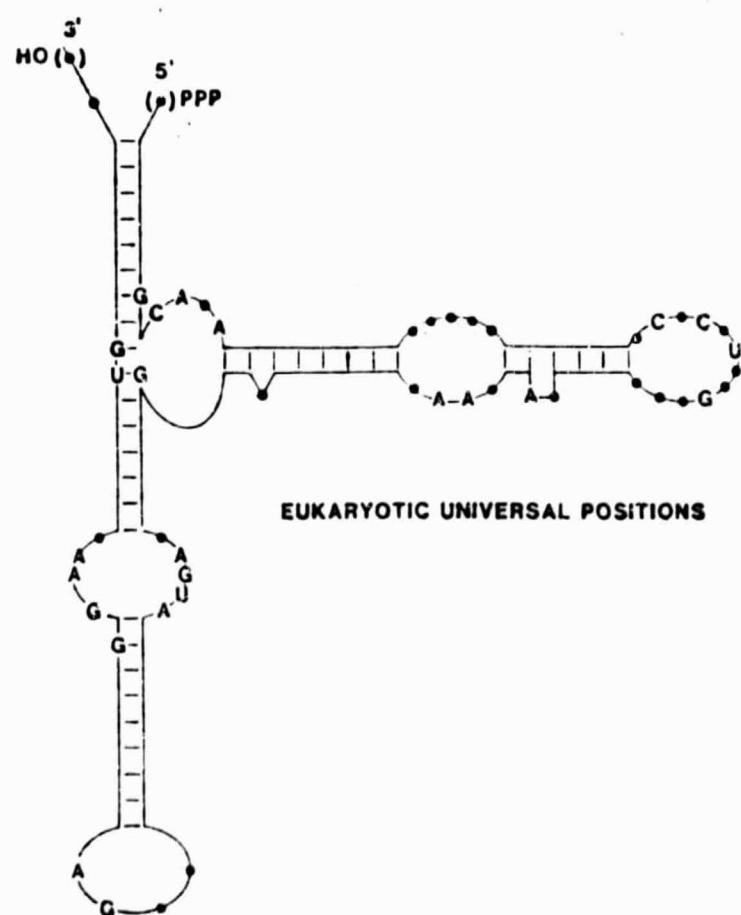
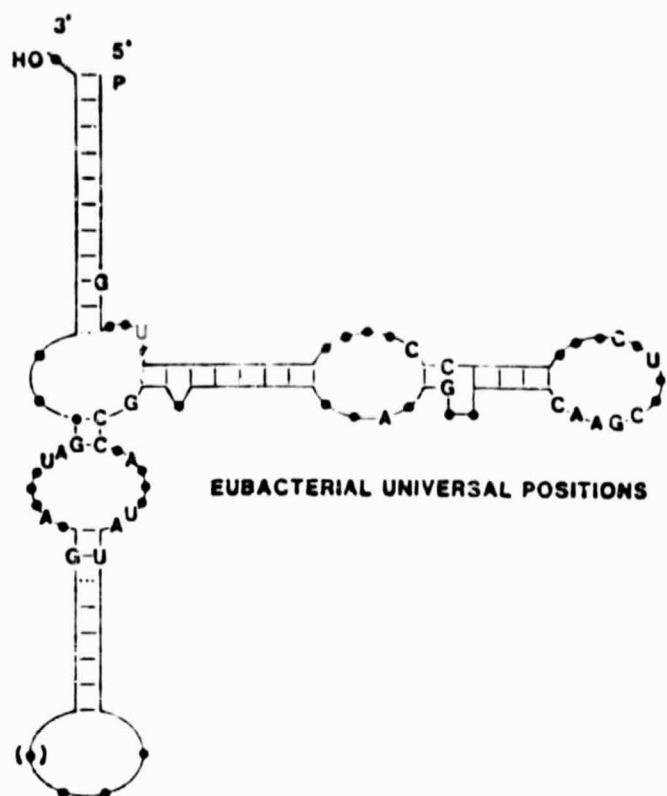


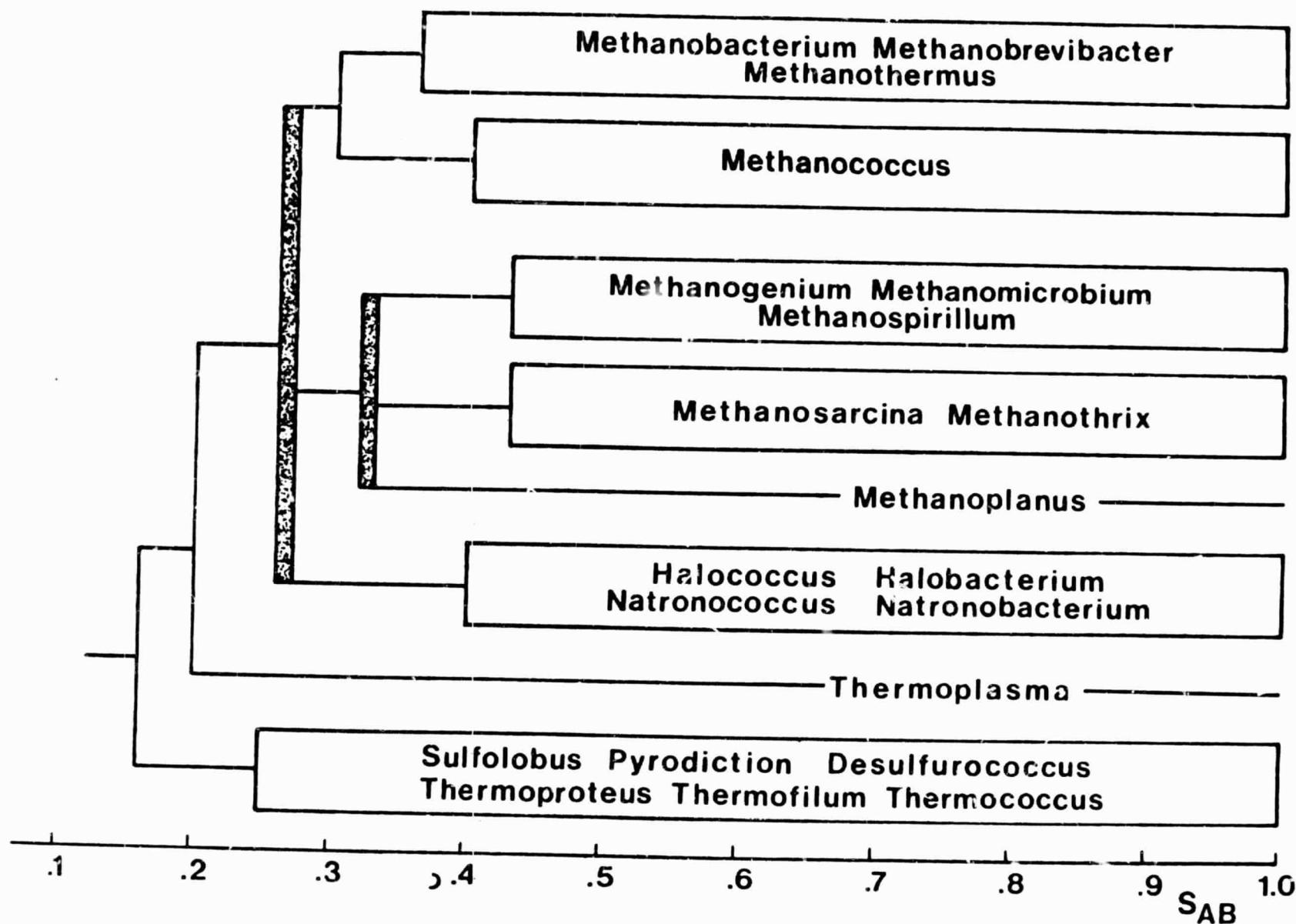
Figure 4: Universal positions in 5S rRNA.

degenerated to complete sequence data) exists. As before the data reveals two major groups, the archaeabacteria and the eubacteria. The archaeabacteria, Figure 5, contain at least two divisions and perhaps three. The first of these includes all the methanogens and the extreme halophiles. The second includes Sulfolobus and many thermoacidophiles which have now been found to be related to Sulfolobus. In spite of the increased number of archaeabacteria in the data base, Thermoplasma remains an enigma with no specific relatives. It can be regarded as a representative of a third division or as a very distant member of the methanogen/halophile division.

The most important findings recently have been among the eubacteria. A summary dendrogram shows at least nine divisions. Three of these were unknown in 1980, the Bacteroides/Cytophaga group, a mixed cluster of sulfur dependent bacteria, Myxococci and Baellivorios, and the Planctomyces/Pirella group. Other major divisions have been further clarified and include Chlorobium and its relatives, Chloroflexus and its relatives, the Spirochaeta which may include Leptospira and Haloanerobium, the purple bacteria, the traditional gram positive bacteria, the cyanobacteria and the isolated cluster Deinococcus.

Several representative figures are shown here to illustrate the results. Figure 6 is a schematic diagram based on S_{AB} values which shows the structure of the very extensively studied purple group. The three divisions, alpha, beta and gamma are well separated and each contains photosynthetic branches (shaded) which supports the view that photosynthesis may have been an aboriginal trait in this group. Dendograms such as Figure 7, which displays the detailed results for the enterics and their most immediate relatives, are available for all the strains. Figure 8 is a schematic summary of the results among the high G-C gram

Figure 5: Archaeabacterial phylogeny from 16S rRNA data.



alpha

gamma

beta

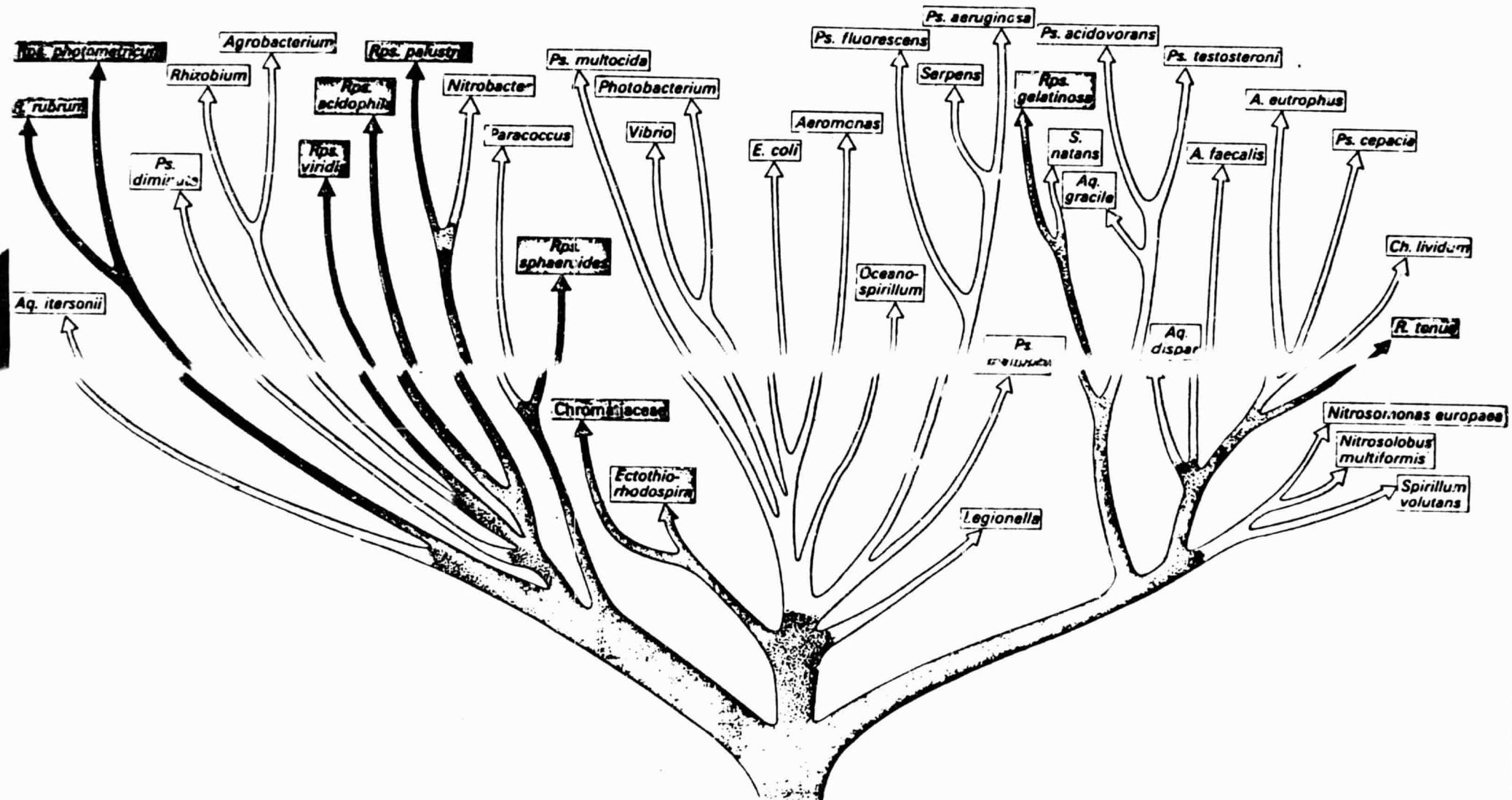
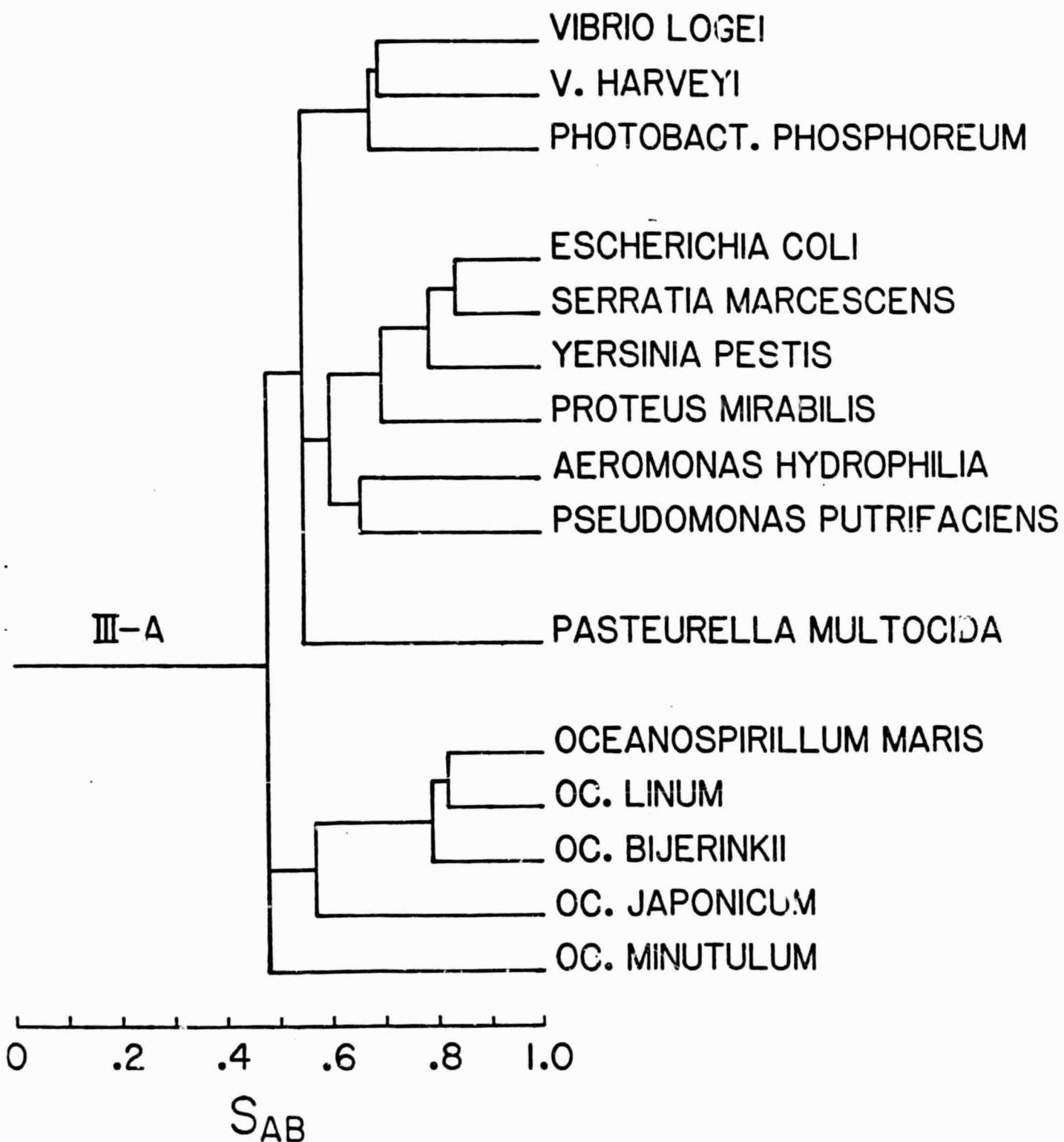


Figure 6: Phylogeny of the “purple bacteria” derived from 16S rRNA data.

Figure 7



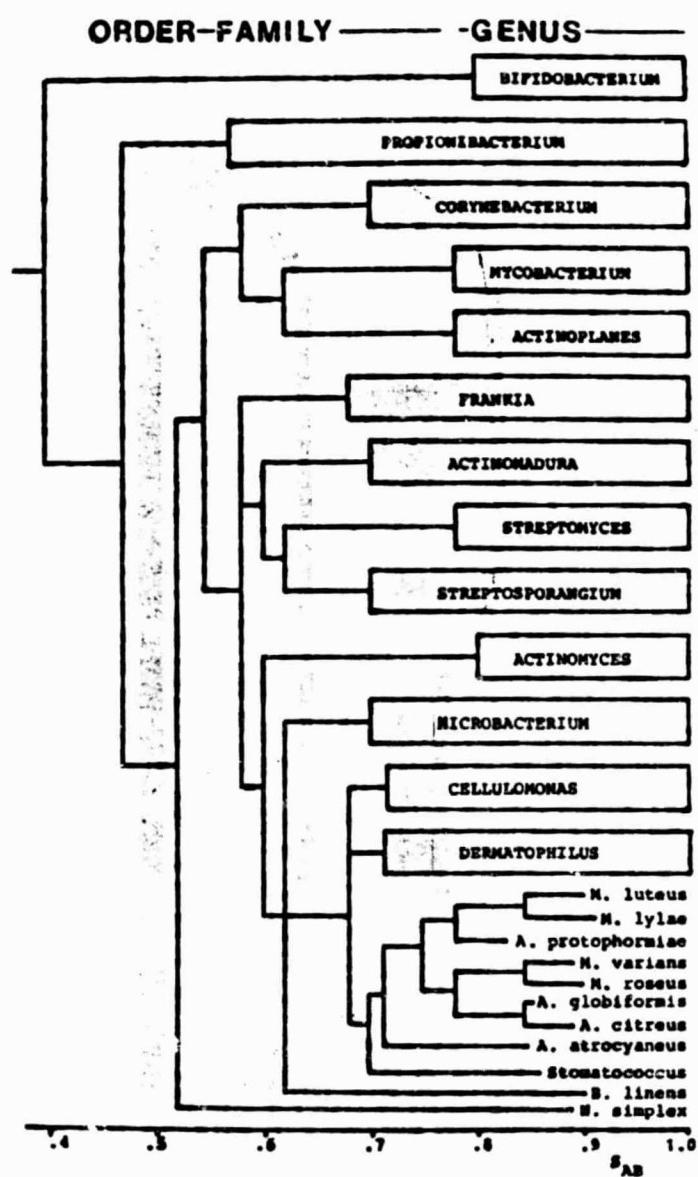


Figure 9: Phylogenetic outline for high %G-C gram positive bacteria from 16S rRNA data.

positive bacteria and also serves to illustrate how a flexible definition of taxonomic terms allows one to integrate the 16S rRNA results into microbial systematics without totally disrupting existing nomenclature. This subject has concerned Professor Stackebrandt and I, and it is discussed in detail in a forthcoming review. It is not useful to include countless dendograms here, but it should be noted that these will now be made available to interested investigators upon request and will appear in a variety of forthcoming publications.

C. Related 16S rRNA Oligonucleotides

Identification of related oligonucleotide fragments is desired in order to form the basis of more sophisticated methodology for comparing catalogs and to fully utilize the catalog data for evaluating in a comparative fashion the validity of secondary and tertiary structural constraints. During the past year our efforts in this area have continued and the major outlines of this family analysis are taking shape. Fifty five major families have been identified and many of these span essentially the entire data set of 450 catalogs. A list of members is augmented by a matrix which specifies the family member actually present in each organism. In the past period we have greatly improved the program OLIGO which is extensively used in generating the families and have added a new program that allows us to change the key names of each catalog so that the matrices can be readily related to phylogenetic relationships established by our clustering programs.

D. Oligonucleotide Maps

A major improvement in the past period was the development of an oligonucleotide dictionary which indicates where every oligonucleotide is found.

Subsequently this dictionary is being used to determine signature properties of each oligonucleotide. This is accomplished by a mapping program which indicates on a dendrogram, all the organisms which contain a particular oligomer, see Figure 9. These maps reveal oligomers that are highly characteristic of small groups or major clusters. A few oligomers are only found in isolated organisms, and any coincidences involving them are clearly random. One also sees oligomers which are found in two or more widely divergent clusters so that one can suspect that their coincidence is the result of evolutionary convergence.

In principle this information can be used to refine trees based on 16S rRNA catalog data. In practice the improvement so obtained probably would not justify the time and effort required. These oligomer maps are useful in testing the reasonableness of related oligomers (see previous section), and defining oligonucleotide signatures that can be used for identifying organisms in mixed culture systems or ecological settings. By expanding the individual signatures of each oligomer into group signatures, one can then use them to test the position of individual organisms on the dendrogram. This type of refinement of the SAB picture works well with "fast clock" organisms that retain the signature of a particular grouping but do not cluster with it due to excessive changes in the non-characteristic oligomers.

III. OTHER RESULTS

A. rRNA Data Base and Software

To support our activities we have sought to assemble a data base of all rRNA complete sequence data. This task is largely accomplished though some checking of sequences is still required. We have also developed software for working with this data.

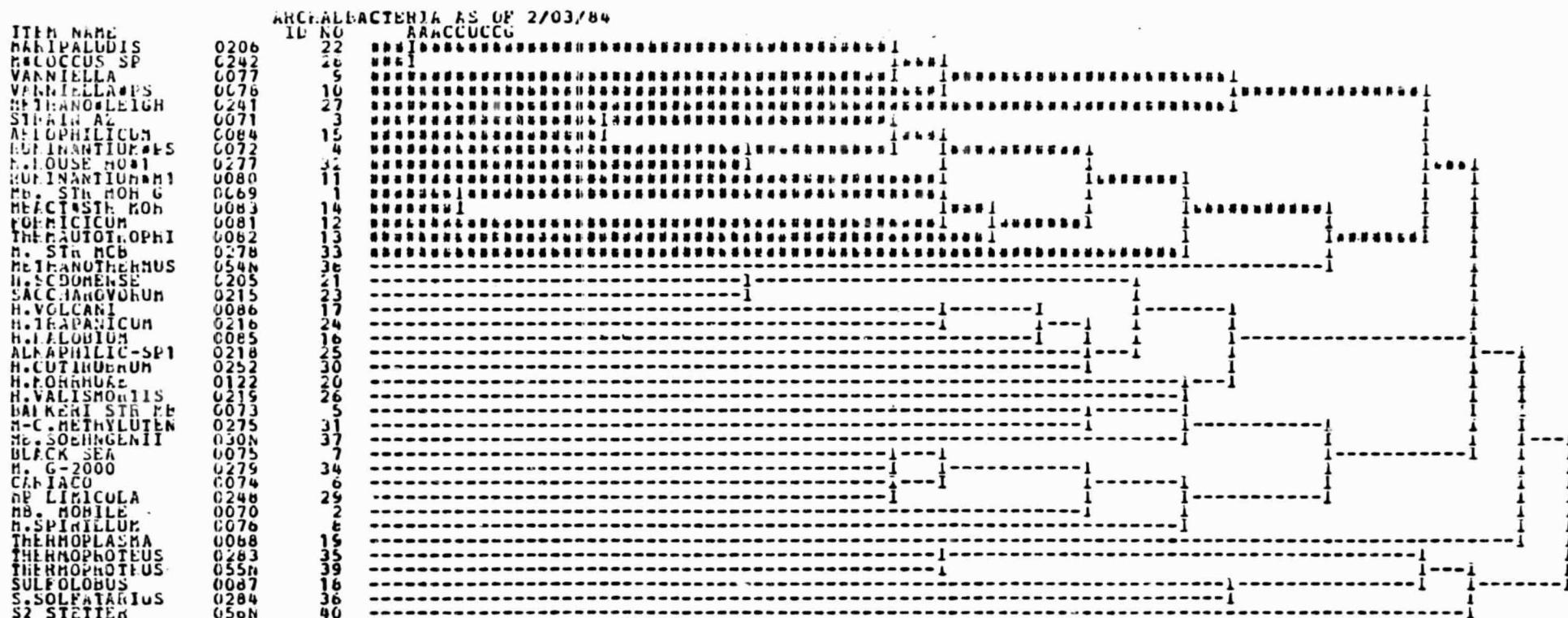


Figure 9: Occurrence of AAACCUCCG in the archaeabacteria.

The information problem in this field at present cannot be underestimated. To deal with it, we use a master sheet approach which allows us to rapidly display what is known or has been proposed. For example we have a separate printout for each of the proposed secondary structure models for L rRNA that have been proposed by others.

The concept is illustrated in Figure 10. Here a small section of the large subunit, rRNA is displayed. The likely helical regions are in brackets and the program tests the individual base pairs to see if they really pair as proposed. If they do, the letter is printed in blue, if not it is red. Yellow letters indicate wobble pairs. In the area outside the helices all the bases are indicated in black except those from Hb. volcani which differ from E. coli. This illustrates how we highlight the differences between sequences. We can for example compare all the sequences to a consensus sequence in some or all areas. Other software utilizes a base by base coloring scheme that is excellent for sequence alignment.

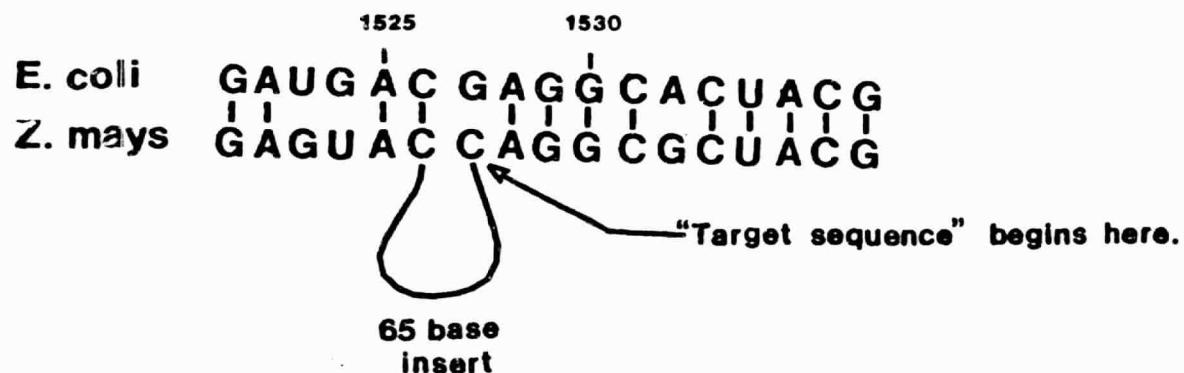
B. unusual 23S rRNA Insertion

While examining potential secondary structures in 23S rRNA we serendipitously discovered that a large insertion in maize chloroplast 23S rRNA relative to E. coli 23S rRNA apparently involves a partial duplication phenomenon, Figure 11. This, coupled with our earlier discovery of a large insertion in Halococcus morrhuae 5S rRNA, has led us to the realization that such large insertion/deletion events are crucial in RNA evolution. Further examination of all such inserts is needed to determine if an obvious common pattern is found. Certainly the duplication seen here along with the very extensive inverted

RAT0.28 UGUUUU-AAUUA-AAACA-AA-[(GCAUCGCG) -AAGG-[(CCC) -GCGGC-(GGG)] UGUUGA-(CGCGUGU)] G
RAT1.28 UGUUUU-AAUUA-AAACA-AA-[(GCAUCGCG) -AAGG-[(CCC) -GCGGC-(GGG)] UGUUGA-(CGCGAUGU)] G
MOU.28 UGUUUU-AAUUA-AAACA-AA-[(GCAUCGCG) -AAGG-[(CCC) -GCGGC-(GGG)] UGUUGA-(CGCGAUGU)] G
XE.28 UGUUUU-AAUUA-AAACA-AA-[(GCAUCGCG) -AAGG-[(CCC) -GAGGC-(GGG)] UGUUGA-(CGCGAUGU)] G
PHY.26 UGUUAU-AAUUA-AAACA-UA-[(GCGAUUUG) -UUGG-[(UGC) -CAAA. (GCU)] UGUAAA-(CIAAUUGU)] G
YEA.26 UGUCU-AAUUA-AAACA-UA-[(GCAUUGCG) -AUGG-[(UCA) -GAAAG (UGA)] UGUUGA-(CGCAAUGU)] G
YEB.25 UGUCU-AAUUA-AAACA-UA-[(GCAUUGCG) -AUGG-[(UCA) -GAAAG (UGA)] UGUUGA-(CGCAAUGU)] G
H.VOL UGUCU AGUAA CAACA UA [(GUGACCG) CAAA [(UCC) GCAA. (GGA)] CUCGUA (CGGUCACU)] G
BACC. STE UGUUU AUCAA AAACA CA [(GGUCUCU) CGAA [(GUC) GUAA. (GUC)] GACGUA-(AGGGCU)] G
E.COLI UGUUU-AUUA-AAACA-CA-[(GACUCGU) -CAAA-[(CAC) -GAAA. (GUG)] GACGUA-(ACUGUGU)] G
MZE.CHL UGUAU ACCAA AAACA CA [(GGUCUCU) CAAA [(GUC) GUAA. (GAC)] CAUGUA (GGGGCU)] G
TOB.23 UGUUU-ACCAA-AAACA-CA-[(GUCUCU) -CAAA-[(GUC) -GUAA. (GAC)] CAUGUA-(GGGGCU)] G
CYA.BE UGUUU-ACCAA-AAACA-CA-[(GUCUCU) -CUAA-[(GUC) -GUAA. (GAC)] GAUGUA (GGGGCU)] G
MZE.MIT UGUUU AUUA AAACA CA [(GACUCU) CUAA [(GUG) GUAA. (CAC)] GAUGUA (AGAGUCU)] G

FIGURE 10: LOCAL STRUCTURAL REGION IN 23S rRNA.
Base changes between *Hb. volcani*
and *E. coli* are highlighted in the
putative single stranded regions.

A. Location of insertion:



B. Homology suggests duplication at insertion site:

Target CAGGCGCUACGGCGCGAA · GUAUG
5' portion CAAGCGCUGCAGCGCGAAAGUAUG
of insert

Figure 11: Possible duplication associated with a 23S rRNA insertion.

repeat pattern of the unduplicated portion of the maize insert is very intriguing. We are currently assembling information about all known insertion/deletion events in the rRNAs.

C. Archaeabacterial 7S RNA

While conducting the 5S rRNA work on the extreme halophiles we consistently encountered a novel RNA in large amounts. This RNA was found to be a homogeneous RNA species by RNA fingerprinting, Figure 11 and Table 3, containing by gel mobility studies 325-375 nucleotides. A similar RNA has been seen in all archaeabacteria examined. This 7S RNA is not, as far as we can tell from existing data, homologous to any rRNA, and it is not found in purified 70S ribosomes. Further work with this molecule is being supported by the new NSF grant.

IV. DISSEMINATION OF RESULTS

(A) Presentations

It is the policy of the principal investigator to expedite the dissemination of the results of this study to the public domain. In order to facilitate this process, a variety of presentations have been made in the previous funding period. These include the following invited presentations:

- (1) Invited seminar speaker at Department of Biochemistry, Dalhousie University, Halifax, Nova Scotia, Canada, September 18, 1981.
- (2) Invited symposium presentation, Annual Meeting Canadian Society of Microbiologists, Quebec City, Canada, June 9, 1982.
- (3) Speaker at First Symposium on Chemical Evolution and the Origin and Evolution of Life at NASA AMES Research Center, Moffett Field, California, August 3, 1982.
- (4) Invited symposium speaker at XIII International Congress of

Hb. halobium
NCMB 777

Figure 12: Hb. halobium 7S RNA Fingerprint.

TABLE 3

Oligonucleotide Catalog of Hb. volcanii 7S RNA with an Indication of which Oligomers are Present or Absent in Hb. vallismortis and Hb. halobium.

<u>Hb. volcanii</u>	<u>Hb. vallismortis</u>	<u>Hb. halobium</u>
G	+	+
CG	+	+
CCG	+	+
CAG	+	+
ACG	+	+
AAG	+	+
CCAG	+	-
ACCG	(AC,C)G	(AC,C)G
ACAG	-	-
AACG	+	-
CCCCCG	-	-
CCAACG	-	-
CAACCG	-	-
AACACG	(AAC,AC,C?)G	(AAC,AC,C?)G
UG	+	+
CUG	-	+
UCG	+	+
UAG	+	-
UCCG	+	(U,C,C)G
CUCG	+	(C,U,C)G
ACUAG	-	(AC,U)AG
AAAUAG	-	-
UACCCCG	(U,AC,C)G	-
ACCAUCG	-	-
(AAAC,C)UG	-	-
CAACCUCG	(AAC,C,U)G	-
UUCG	?	-
UUAG	-	+
UUCCG	-	-
UUUAG	?	-
CUUUCG	-	-
UUCAUCCG	-	-
terminii:		
pUUUG	?	?
CA _n CCUX _n OH	?	?
2 undetermined		

Microbiology, Boston, Massachusetts, August 11, 1982.

(5) Invited seminar speaker at Department of Botany at University of Texas, Austin, Texas, October 1982.

(6) Invited seminar speaker at Department of Microbiology, State University of New York at Stony Brook, December 3. 1982.

(7) Invited seminar speaker at the Department of Bacteriology, University of California at Davis, April 31, 1983.

(8) Invited symposium speaker at conference entitled "Evolution: Shaping of Molecules, Microbes and Complex Organisms", University of Colorado, Boulder, Colorado, May 1, 1983.

(9) Invited symposium speaker at the 23rd Annual Meeting American Society for Cell Biology in San Antonio, Texas, December 1, 1983.

(10) Invited seminar speaker at the Department of Microbiology at Ohio State University, Feb. 23, 1984.

(11) Invited seminar speaker at The Biological Laboratories at Harvard University, April 26, 1984.

(12) Invited instructor at NASA sponsored summer course in Planetary Biology and Microbial Ecology at San Jose State University, July 11-17, 1984.

(13) Invited seminar speaker at the Department of Biochemistry at Stockholm University in Stockholm, Sweden, September 28, 1984.

(14) Invited seminar speaker at NASA-JSC Microgravity Science Seminar, October 25, 1984.

In the forthcoming months the principal investigator will be making presentations at the annual American Society of Microbiology meeting in Las Vegas, the EMBO workshop on the Molecular Genetics of Archaebacteria in Munich and the workshop on the Biology and Biochemistry of Archaebacteria in Martinsried.

(B) Publications:

Publications represent the final step in dissemination of knowledge into the public sphere. Since 1982 the following manuscripts directly attributable to the studies supported by this NASA grant, NSG-7440 appeared. These are:

1. Fox, G. E. and Woese, C. R., "Archaeabacterial 5S Ribosomal RNA" Zbl. Bakteriol. Hyg. I. Abt. Orig. C3, 330-345 (1982).
2. Tanner, R.S., Stackebrandt, E., Fox, G.E., Gupta, R., Magrum, L.J. and Woese, C. R. "A Phylogenetic Analysis and Anaerobic Eubacteria Capable of Synthesizing Acetate from CO₂" Current Microbiol.: 7: 127-132 (1982).
3. Nicholson, D. E. and Fox, G. E. "Molecular Evidence for a Close Phylogenetic Relationship Between Box-Shaped Halophilic Bacteria, Halobacterium vallismortis and Halobacterium marismortui" Can. J. Microbiol. 29, 52-59 (1983).
4. Sobieski, J., Chen, K. N., Filieatreau, J., Pickett, M. H. and Fox, G.E. "16S rRNA Oligonucleotide Catalog Data Base" Nucl. Acids Res., 12, 141-148 (1984).
5. Deming, J. W., Hada, H., Colwell, R. R., Leuhrsen, K. R. and Fox, G. E. "The Ribonucleotide Sequence of 5S rRNA from Two Status of Deep-Sea halophilic Bacteria" J. Gen. Microbiol. 130, 1911-1920 (1984).

Manuscripts which are currently in press are as follows:

6. Woese, C. R., Stackebrandt, E., Weisburg, W., Pasteur, B. J., Madigan, M. T., Fowler, V. J., Hahn, C.M., Blanz, P., Gupta, R., Nealson, K. H. and Fox, G. E. "The Phylogeny of Purple Bacteria: The Alpha Subdivision", Syst. Appl. Micro., in press.
7. Fox, G. E. "The Structure and Evolution of Archaeabacterial RNA", in The Bacteria Vol. 8, (I.C. Gunsalus, ed.). Academic Press, New York, in

press.

8. Stackebrandt, E., Ludwig, I. and Fox, G. E. "16S Ribosomal RNA Oligonucleotide Cataloging" Methods in Microbiol. Vol. 19, in press.
9. Daniels, C., Hoffman, J. D., Leuhrsen, K. R., Woese, C. R., Fox, G. E., and Doolittle, W. F. "Sequence of the 5S rDNA cistrons of Halobacterium volcani" Mol. Gen. Genetics, in press.
10. Ludwig, W., Seewaldt, E., Kilpper-Buz, R., Schleifer, K.-H., Woese, C.R., Fox, G. E. and Stackebrandt, E. "The Phylogenetic Position of Streptococcus and Enterococcus" J. Gen. Microbiol., in press.

Three additional manuscripts have recently been submitted for publication.

These are:

1. Fox, G.E. and Stackebrandt, E., "The Application of 16S rRNA Cataloging and 5S RNA Sequencing in Microbial Systematics" submitted by invitation to Methods in Microbiology, Vol. 18.
2. Leuhrsen, K. R., Nicholson, D. E. and Fox, G. E., "Widespread Distribution of a Non-Ribosomal 7S RNA in Archaeabacteria" submitted to J. Biol. Chem.
3. Delihauz, N., Leuhrsen, K. R., Gibson, J. and Fox, G. E., "5S rRNA Sequences from the Purple Photosynthetic Bacteria" submitted to FEBS Let.

Further papers will be forthcoming in the next 12 months. These will include several reports on the 5S rRNA and 16S rRNA results described herein and a theoretical paper on the importance of knowledge of the translation process in molecular evolution.

(C). Abstracts

In addition to the presentations and papers described in the previous two sections, two abstracts were published in the past grant period that reported results obtained under the aegis of NSG-7440. These were:

1. Luehrsen, K. R., Eubanks, D. C., Fox, G. E. and Friz, C. T.

"Phylogenetic Relationships Between Three Species of Free Living Amoeba" J. Protozoology, 29:498-499 (1982).

2. Deming, J. W., Luehrsen, K. R., Fox, G. E., Hada, H. and Colwell, R. R.

"Phylogenetic Position and Two Strains and Deep Sea Bacteria" Abstracts 3rd International Conference on Microbial Ecology, East Lansing Michigan (1983).